233 Analyse numérique Matricielle. Résolution approchée de systèmes linéaires, recherche d'éléments propres, exemples.

Ref : [F.C] F : [illegible] Analyse numérique [Qua] Quarteroni : Méthodes numériques pour le calcul scientifique.
[Cia] Ciarlet, Analyse numérique matricielle et optimisation

Dev :
19 Convergence méthode itérative (+ Householder) [Cia]
28 Gradient à pas optimal.

Cadre : On se place dans $K = \mathbb{R}$ ou $\mathbb{C}$, pour $n \in \mathbb{N}^*$

I. Normes matricielles et conditionnement.

1) Normes matricielles

[F.C] 9

Def 1 : On appelle norme matricielle une norme sur l'espace vectoriel $\mathcal{M}_n(K)$. On dit qu'une norme $\|.\|$ est sous multiplicative si $\forall A, B \in \mathcal{M}_n(K)$, $\|AB\| \leq \|A\| \, \|B\|$.

Def 2 : Pour $|.|$ une norme sur K^n, l'application $\|.\|$ définie sur $\mathcal{M}_n(K)$ par $\|A\| = \sup_{x \neq 0} \frac{|Ax|}{|x|} = \sup_{|x|=1} |Ax|$ définit une norme sur $\mathcal{M}_n(K)$, dite norme matricielle subordonnée à la norme $|.|$.

[Qua] 10 24

Prop 3 : Une norme matricielle subordonnée (NMS) est une norme matricielle sous multiplicative.

Ex 4 : Si $|.| = \|.\|_1$ sur $\mathbb{R}^n$, $\|A\| = \sup_{j=1..n} \sum_{i=1}^{n} |a_{ij}|$.

Ex 5 : Toutes les normes matricielles ne sont pas sous multiplicatives : Si $\|(a_{ij})\| = \sup |a_{ij}|$, alors pour $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\|A^2\| = 2 > \|A\| \, \|A\| = 1$.

En particulier, il existe des normes non subordonnées à une norme vectorielle.

[Cia]

Def Prop 6 : Pour $A \in \mathcal{M}_n(K)$ on définit $\rho(A)$ le rayon spectral de A comme le supremum des modules de ses valeurs propres. La norme subordonnée de la norme euclidienne (hermitienne) est alors donnée par $\|A\| = \sqrt{\rho(A^* A)}$ où A^* est l'adjoint de A. ⊛

2) Conditionnement d'une matrice.

[F.C]

Def 7 : Soit $A \in Gl_n(K)$, on définit le conditionnement de A relativement à une norme subordonnée $\|.\|$ le nombre $K(A) = \|A\| \, \|A^{-1}\|$.

Le conditionnement permet de mesurer la sensibilité de la solution d'un système linéaire à la condition initiale :

Si x est solution de $Ax = b$, et $x + \delta x$ est solution de $A(x + \delta x) = b + \delta b$ alors $\frac{|\delta x|}{|x|} \leq K(A) \frac{|\delta b|}{|b|}$. Si $K(A)$ est grand, l'erreur relative en b peut induire une grande erreur en x. On parlera de système mal conditionné.

⊛ Théo 8 (Householder) Si $\|.\|$ est une NMS, alors $\|A\| \geq \rho(A)$, et pour $\varepsilon > 0$, $\exists \|.\|_{\varepsilon, A}$ telle que $\|A\|_{A,\varepsilon} \leq \rho(A) + \varepsilon$

Ex 9 : La matrice de Hilbert est mal conditionnée.

[F;C]

Prop 10 : Pour $A \in Gl_n(K)$, on a $K(A) = K(A^{-1}) \geq 1$, $K(I_n) = 1$ et $K(\lambda A) = K(A)$

Rq 11 : On ne peut pas améliorer le conditionnement par une simple multiplication scalaire. Il faut multiplier à gauche par une matrice bien choisie. Idéalement A^{-1}, mais bien sûr en pratique, on doit être plus subtil.

Prop 12 : Si A est hermitienne, alors $\|A\|_2 = \rho(A)$. De plus si A est inversible et $\lambda_1 \dots \lambda_n$ ses vp rangées par ordre croissant, on a $K(A) = \lambda_n / \lambda_1$.

II. Méthodes directes de résolution des systèmes linéaires.

1) Méthode de Gauss.

[Qua]

On se ramène dans tous les cas à l'étude de systèmes triangulaires "rapides" à résoudre par descente et remontée.

Pour un système $Ax = b$ où $A = (a_{ij})$, on considère les opérations suivantes.

- → Choisir une ligne de A telle que $a_{i1} \neq 0$.
- → Echanger la ligne en question avec la première.
- → Remplacer les lignes $2 \to n$ par $l_i - \frac{a_{i1}}{a_{11}} l_1$
- → Répéter sur la sous-matrice $(a_{ij})_{i,j \geq 2}$.

On termine cet algorithme avec un système triangulaire supérieur. Mais on peut avoir des problèmes de nullité des pivots.

2) Factorisation LU.

Théo 13 : Soit $A \in Gl_n(K)$ telle que toutes les sous matrices principales de A soient inversibles, il existe un unique couple (L, U) constitué d'une matrice triangulaire inf à diagonale unité, et d'une matrice triangulaire supérieure, tel que $A = LU$.

Cor 14 : $A \in S_n^{++}(\mathbb{R})$ admet une décomposition LU.

Ex 15 : Si A est tridiagonale $a_2, \dots a_n$, $b_1 \dots b_n$, $c_1 \dots c_{n-1}$, on définit la suite $\delta_0 = 1$, $\delta_1 = b_1$, $\delta_{k+2} = b_{k+2}\delta_{k+1} - a_{k+2} c_{k+1} \delta_k$, $k \geq 0$.

Si tous les δ_k sont différents de 0, la factorisation LU de A est

$$\begin{pmatrix} 1 & & 0 \\ a_2 \frac{\delta_0}{\delta_1} & 1 & \\ & a_n \frac{\delta_{n-2}}{\delta_{n-1}} & 1 \end{pmatrix} \begin{pmatrix} \delta_1/\delta_0 & c_1 & 0 \\ & \delta_2/\delta_1 & c_{n-1} \\ 0 & & \delta_n/\delta_{n-1} \end{pmatrix}$$

Ex 16 : Résolution de la matrice du Laplacien discret.

Méthode de calcul d'une décomposition LU : Méthode de Doolittle.

Pour $k = 1 : n$ poser $u_{kj} = a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj}$ $\quad l_{ik} = \frac{1}{u_{kk}}\left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk}\right)$

[Fil] 2729.

3) Factorisation de Cholesky

Théo 17 : Soit $A \in \mathcal{M}_n(K)$ hermitienne (symétrique) définie positive. Alors il existe une matrice R triangulaire supérieure telle que $R^* R = A$. Si on impose les coefficients diagonaux de R comme positifs, alors la factorisation est unique.

Méthode de calcul. $h_{11} = \sqrt{a_{11}}$, pour $i = 2 : n$, faire
$h_{ij} = \frac{1}{h_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} h_{ik} h_{jk}\right)$ $j \in [\![1, i-1]\!]$. $- h_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} h_{ik}^2\right)^{1/2}$

on obtient ainsi R^*.

Ex 18 : $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ (-\sqrt{2})^{-1} & \sqrt{3}/\sqrt{2} & 0 \\ 2\sqrt{2}/\sqrt{3} & \sqrt{2} - \frac{\sqrt{3}}{2} & \sqrt{2\sqrt{3} - \frac{25}{6}} \end{pmatrix}$

III. Méthodes itératives de résolutions des systèmes linéaires.

On cherche étant donné un système linéaire $Ax = b$, à construire une suite x^n de $\mathbb{R}^n$ qui converge vers la solution. On peut relier ce problème à un problème de minimisation. Si $A \in S_n^{++}(\mathbb{R})$, la fonctionnelle quadratique $J(x) = \frac{1}{2}(Ax, x) - (b, x)$ admet un unique minimum qui se trouve être la solution du système. Les méthodes que nous allons exposer ont des analogues en optimisation.

1) Généralités.

[Cia] 97-100

On se ramène à résoudre $u = Bu + c$ où $(I_n - B) \in GL_n(K)$ est telle que la solution de ce dernier système soit également la solution de $Au = b$. On est alors tenté d'appliquer un algorithme de point fixe et de poser

$$u_{n+1} = Bu_n + C. \quad (M)$$

la liberté résidant dans le choix de la matrice B et du vecteur C.

Théo 19 : On a équivalence entre :
i) la méthode (M) converge ii) $\rho(B) < 1$ iii) $\|B\| < 1$ pour au moins une norme matricielle subordonnée $\|.\|$. DVD

Rq 20 : La convergence est géométrique et indiquée par $\rho(B)$ (ou pire).

2) Méthodes itératives classiques.

On suppose ici pouvoir écrire A sous la forme $M - N$ où $M \in GL_n(K)$ est 'facile' et inverser (en pratique, M diagonale ou triangulaire). On a alors

$Au = b \Leftrightarrow Mu = Nu + b \Leftrightarrow u = M^{-1}(Nu + b) = M^{-1}Nu + M^{-1}b$. D'où la méthode $u_{k+1} = M^{-1}Nu_k + M^{-1}b$, correspondant à $B = M^{-1}N = I - M^{-1}A$. On remarque d'emblée que $I - B = M^{-1}A$ est inversible.

[Cia] 97/106

Notation : Pour $A \in GL_n(K)$, on pose D la partie diagonale, $-E$ et F ses parties triangulaires inférieures et supérieures strictes (respectivement) de sorte que $A = D - E - F$

Méthode de Jacobi ; on pose $M = D$ (donc $N = E + F$) on obtient la méthode itérative $Du_{k+1} = (E + F)u_k$, la matrice correspondante est $J = D^{-1}(E + F)$.

Méthode de Gauss Seidel On prend $M = D - E$, donc on travaille avec la matrice $\mathcal{L}_1 = (D - E)^{-1}F$, dite matrice de Gauss Seidel.

Méthode de relaxation : On introduit un paramètre ω, dit de relaxation et on pose $M = \frac{D}{\omega} - E$, on a donc la matrice d'itération $\mathcal{L}_\omega = \left(\frac{D}{\omega} - E\right)^{-1}\left(\frac{1-\omega}{\omega}D + F\right)$.
on pourra alors chercher le meilleur paramètre ω pour minimiser $\rho(\mathcal{L}_\omega)$.

Théo 21 : Si A est hermitienne définie positive, $A = M - N$ où $M \in GL_n(K)$. Alors si $M^* + N$ est définie positive, on a $\rho(M^{-1}N) < 1$

Théo 22 : Si A est hermitienne définie positive, la méthode de relaxation converge pour $\omega \in]0, 2[$ (en particulier, Gauss Seidel converge).

Théo 23 : On a l'inégalité $\rho(\mathcal{L}_\omega) \geq |\omega - 1|$, $\omega \neq 0$. Ainsi la méthode de relaxation ne peut converger que pour $\omega \in]0, 2[$

Théo 24 : Si A est tridiagonale par blocs. On a $\rho(\mathcal{L}_1) = \rho(J)^2$: les méthodes de Jacobi et de Gauss Seidel convergent donc simultanément, et si elles convergent, la méthode de Gauss Seidel converge plus rapidement.
Si A est de plus hermitienne définie positive. Alors le paramètre de relaxation optimal est $\omega = 2/(1 + \sqrt{1 - \rho(J)^2})$ si $\rho(J) > 0$, et 1 si $\rho(J) = 0$

3) Méthodes d'optimisation : Cas réel.

[Cia] 186

On considère $J : x \mapsto J(x) \in \mathbb{R}$ (moralement une fonctionnelle quadratique), supposément une fonctionnelle α-convexe (matrice $S^{++}(\mathbb{R})$). On construit des méthodes itératives sous la forme $x^{n+1} = x^n + \rho^n d^n$ (ρ^n est le pas, d^n la direction).

Méthode de relaxation. On prend successivement d^n les vecteurs de base canonique : Étant donné $x^n \in \mathbb{R}^d$ on pose, pour $i \in [\![1, n]\!]$ $x_i^{n+1} \in \mathbb{R}$ minimisant la fonction $t \mapsto J(x_1^{n+1}, \dots, x_{i-1}^{n+1}, t, x_{i+1}^n, \dots, x_d^n)$, on pose ensuite $x^{n+1} = (x_1^{n+1}, \dots, x_d^{n+1})$

Théo 25 : Si J est une fonctionnelle elliptique (A est définie positive). Alors la méthode de relaxation converge.

Rq 26 : C'est en fait la méthode de Gauss Seidel appliquée au système linéaire correspondant.

[Cia] 189 192.

On peut également prendre $-\nabla J(x^n)$ comme direction, ce qui donne les méthodes de gradient dont la liberté réside dans le choix du pas.

Gradient à pas fixe : $\rho = \rho^n$ est constant.

Gradient à pas optimal : ρ^n est choisi pour minimiser $\rho \mapsto J(x^n - \rho \nabla J(x^n))$.

Théo 27 : Si J est α-convexe, et ∇J est L-lipschitzien, alors pour $\rho \in]0, \frac{2\alpha}{L^2}[$, la méthode de gradient à pas fixe converge.

Théo 28 : Si J est α-convexe, alors la méthode du gradient à pas optimal converge. DVP

Rq 29 : L'encadrement du théorème 27 n'est pas optimal, pour une fonctionnelle quadratique on peut aller jusqu'à $2/\lambda_n$ au lieu de $2\lambda_1/\lambda_n^2$ (le pas fixe optimal est $2/(\lambda_1+\lambda_n)$).

Rq 30 : Le pas optimal peut en général être coûteux à calculer, mais pour une fonctionnelle quadratique, il est explicite et donné par $\|r^k\|^2/(Ar^k, r^k)$ où $r^k = Ax^k - b$.

IV. Problèmes aux valeurs propres.

[Qua] 168.

On cherche ici à trouver, étant donnée une matrice A, $\lambda \in \sigma(A)$ et $v \neq 0$ / $Av = \lambda v$.

Théo 31 (Gerschgörin) Soit $A \in \mathcal{M}_n(\mathbb{C})$, posons $A = (a_{ij})_{i,j \in [\![1,n]\!]}$, pour $i \in [\![1,n]\!]$, on pose D_i le disque complexe (fermé) de centre a_{ii} et de rayon $\sum_{j \neq i} |a_{ij}|$, c'est le $i$ème disque de Gerschgörin. On a $\sigma(A) \subseteq \bigcup_{i=1}^{n} D_i$

Rq 32 : Ce résultat donne une première approximation des racines d'un polynôme en fonction de ses coefficients (par les matrices compagnons)

Rq 33 : Une matrice à diagonale strictement dominante est ainsi inversible.

1) Méthode de la puissance.

[Fil] 64 68 [Qua] 172.

On ordonne les valeurs propres de A (avec multiplicité) par leur module : $|\lambda_1| \leq \dots \leq |\lambda_n|$. On suppose $|\lambda_n| > |\lambda_{n-1}|$ (valeur propre dominante). On considère la méthode

$$Ax^k =: y^{k+1} \quad x^{k+1} = y^k/\|y^k\|.$$

La dernière condition évitant la divergence de la méthode.

Théo 34 : Si $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable, $x_0 \notin \mathrm{Ker}(A - \lambda_n I_n)^{\perp}$. Alors la suite (x_n) définie ci-dessus converge vers un vecteur propre associé à la valeur propre λ_n. En outre la suite $(Ax^k)_j$ où $|x^k_j| = \|x^k\|_\infty$ converge vers λ_n.

Rq 35 : En appliquant cette méthode à A^{-1} (inversion de systèmes linéaires) on peut retrouver la plus petite valeur propre de A. Voire même toutes les valeurs propres (puissance inverse sur $(A - \lambda_i Id)$).

2) Méthode de Jacobi.

[Cia] 111 117

Soit $A \in \mathcal{M}_n\mathbb{R}$ symétrique. La méthode de Jacobi consiste à construire une suite (Ω_k) de matrices orthogonales « élémentaires » telle que la suite de $A_{k+1} = {}^t\Omega_k A_k \Omega_k = {}^t(\Omega_1 \dots \Omega_k) A \Omega_1 \dots \Omega_k$, $k \geq 1$ converge vers la matrice $\mathrm{diag}(\lambda_1 \dots \lambda_n)$ où $\{\lambda_i\} = \sigma(A)$ et telle que la suite $O_k = \Omega_1 \dots \Omega_k$ converge vers une matrice orthogonale dont les colonnes forment une base orthonormée de vecteurs propres de A.

Théo 36 : Soit $1 \leq p < q \leq n$ et $\theta \in \mathbb{R}$. On pose

$$\Omega := \begin{array}{c c} & \begin{array}{ccccccc} & & p & & q & & \end{array} \\ & \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \cos\theta & & -\sin\theta & & \\ & & & 1 & & & \\ & & \sin\theta & & \cos\theta & & \\ & & & & & 1 & \\ & & & & & & 1 \end{pmatrix} \begin{array}{c} \\ \\ p \\ \\ q \\ \\ \\ \end{array} \end{array}$$

1) Si A est symétrique, alors $B = {}^t\Omega A \Omega$ est symétrique et $\sum b_{ij}^2 = \sum a_{ij}^2$.

2) Si $a_{pq} \neq 0$ $\exists! \theta \in]-\frac{\pi}{4}, \frac{\pi}{4}] \setminus \{0\}$ / $b_{pq} = 0$. C'est la solution de $\cot(2\theta) = \frac{a_{qq} - a_{pp}}{2a_{pq}}$ et $\sum b_{ii}^2 = \sum a_{ii}^2 + 2a_{pq}$

Lemme 37 Soit X un espace vectoriel normé de dimension finie (x_k) une suite bornée de X admettant un nombre fini de valeurs d'adhérence et telle que $\lim \|x_{k+1} - x_k\| = 0$ converge.

Théo 38 La suite (A_k) définie ci-dessus converge et $\lim A_k = \mathrm{diag}(\lambda_{\sigma(i)})$ où $\sigma \in \mathfrak{S}_n$.

Théo 39 : On suppose que les valeurs propres de A sont distinctes. Alors la suite (O_k) définie ci-dessus converge vers une matrice orthogonale dont les colonnes forment une base orthonormée de vecteurs propres de A.